Environmental Modelling & Software 15 (2000) 79–86

Environmental Modelling & Software

www.elsevier.com/locate/envsoft

Predicting crop yields under climate change conditions from monthly GCM weather projections

V.R. Reddy [a], Ya.A. Pachepsky [b,*]

[a] USDA:ARS:NRI:Remote Sensing and Modeling Laboratory, BARC-West, Bldg. 007, Room 008, Beltsville, MD 20705, USA
[b] Department of Botany, Duke University, Durham, NC 27708, USA

Received 15 November 1998; accepted 25 March 1999

Abstract

Estimation of changes in crop yields is currently based on projections of atmospheric General Circulation Models (GCM) and the use of crop simulators. Crop simulators require daily input of environmental variables. GCMs produce monthly projections of climatic variables. Our objective was to explore the possibility of using monthly weather projections in yield estimates. We considered atmospheric CO_2 level, total solar radiation, average maximum and minimum temperature, and rainfall for five months of the growing season. The group method of data handling (GMDH) was applied to relate crop yields to these variables. Projections of GCMs were downscaled to provide daily weather variables for the Mississippi Delta, and weather patterns were obtained in 50 replications for each GCM. The soybean crop simulator GLYCIM was used to generate crop yields on sandy loam, loam and silt loam soils. The equations built with GMDH explained 81–85% of yield variability, and included solar radiation in July and August, CO_2 level, minimum temperature in June and August, and rainfall in August. Published by Elsevier Science Ltd.

Keywords: Climate change; Group method of data handling; Soybeans; Crop simulation; General circulation models

Software availability

(1) Program title: DWD (*D*ownscaling *w*eather *d*ata)
Developers: Dr Yakov Pachepsky
Contact address: USDA Remote Sensing and Modeling Laboratory, Bldg. 007, Rm. 008, BARC-WEST, Beltsville, MD 20705, USA. Tel.: +1-301-504-7468; fax: +1-301-504-5823; e-mail: ypachepsky@asrr.arsusda.gov
First year available: 1999
Software requirements: Microsoft Windows 95 or higher
Hardware requirements: PC (80486 or better), math coprocessor (recommended), 4Mb hard disk space
Cost: Public domain software

(2) Program title: GPDPTI (Grid Point Interpolation)
Developer: Dennis Joseph
Contact address: Data Support Section, SCD, National Center for Atmospheric Research, PO Box 3000, Boulder, CO 80307, USA. Tel.: +1-303-497-1216; fax: +1-303-497-1804; e-mail: joseph@ucar.edu, datahelp@ucar.edu
First year available: 1993
Software requirements: FORTRAN 77 or later
Hardware requirements: Any PC or UNIX system with FORTRAN compiler
Cost: No charge

(3) Program title: ModelQuest
Developer: AbTech Corporation
Contact address: 1575 State Farm Boulevard, Charlottesville, VA 22911, USA. Tel.: +1-804-977-0686; fax: +1-804-977-9615; e-mail: tech@abtech.com
First year available: 1992
Software requirements: PC, Windows 95
Hardware requirements: 8Mb RAM, math coprocessor, 15 Mb hard disk space
Cost: US$500

(4) Program title: GLYCIM
Developer: Dr Basil Acock
Contact address: USDA Remote Sensing and Modeling Laboratory, Bldg. 007, Rm. 008, BARC-WEST,

* Corresponding author. Tel.: +1-301-504-7468; fax: +1-301-504-5823.
E-mail address: ypachepsky@asrr.arsusda.gov (Y.A. Pachepsky)

1364-8152/00/$ - see front matter. Published by Elsevier Science Ltd.
PII: S1364-8152(99)00011-0

Beltsville, MD 20705, USA. Tel.: +1-301-504-5827; fax: +1-301-504-5823; e-mail: bacock@-asrr.arsusda.gov
First year available: 1991
Software requirements: FORTRAN 77 or later
Hardware requirements: Any PC or UNIX system with FORTRAN compiler
Cost: Public domain software

1. Introduction

Ongoing changes in the global climate are likely to have significant effects on agriculture (Watson et al., 1996). Predicting the direction, scope, and magnitude of these effects is necessary in order to develop and implement successful management responses. Mechanistic crop simulation models are useful tools in this regard because, provided with inputs derived from predicted climate change, they can simulate crop response to these potential future conditions. Atmospheric General Circulation Models (GCM) represent the current state of knowledge about how the earth's climate is likely to behave in the future (Giorgi et al., 1998). GCM projections can be used to provide weather data for future-oriented crop simulations (Rosenzweig and Hillel, 1998). Several comprehensive studies of climate change effects on soybean yields were carried out using crop simulators and GCM projections (Ritchie et al., 1989; Rosenzweig et al., 1994; Haskett et al., 1997; Reddy et al., 1997).

As the ongoing evolution of GCMs themselves produces new climate predictions, the crop simulations should be repeated to give new estimates of crop response. Crop models suitable to simulate elevated CO_2 effects on crop yields include simulation of many complex physiological processes and are thus themselves complex and computationally intensive. Typically, they require daily weather data. GCMs produce monthly projections of weather variables (Giorgi et al., 1998). A simple model explicitly expressing dependence of crop yields on monthly climate variables would be very useful because it could help to characterize these general relationships and, if sufficiently broad in scope, could allow us to estimate crop yields for new GCM projections without recourse to a full run of mechanistic crop simulation.

The complexity of dependencies of crop yields on climate variables lends itself to the use of artificial neural networks (ANN) which are becoming a common tool for modeling complex 'input–output' dependencies (Hecht-Nielsen, 1990). The advantage of ANNs is their ability to mimic the behavior of complex systems by varying the strength of the influence of network components to each other and by varying the structure of the interconnections among components. After establishing network structure and finding coefficients to express the strength of influence of the network components to each other, an artificial neural network becomes a complex formula, relating input values with output values (Alexander and Morton, 1990). This formula can be used like a regression formula.

There is a great variety of ANNs. The back-propagation ANNs are applied most often. These ANNs have been used to reproduce input–output relations for complex process models including agricultural and land use models (Elizondo et al., 1994; Lein, 1995; Joerding et al., 1994; Pachepsky et al., 1996). Although the back propagation ANNs often perform better than conventional statistical regression (Pattie and Haas, 1996), they have some disadvantages compared with regressions. In particular, the equations built during ANN training are opaque, and ANNs do not distinguish inputs by their significance leaving the responsibility to select significant inputs to a user (Jarvis and Stuart, 1996). Recently the group method of data handling (GMDH) has gained popularity as a tool to express complex 'input–output' dependencies. The GMDH provides automated selection of essential input variables and builds hierarchical polynomial regressions of necessary complexity (Farrow, 1984). The networks of polynomials built with GMDH have fewer nodes than the artificial neural networks, but the nodes are more flexible (Hecht-Nielsen, 1990). These networks, called abductive or polynomial networks, have been used in various areas of science and engineering (Sommer et al., 1995; Abdel-Aal and Elhadidi, 1995; Kleinsteuber and Sepehri, 1996) and have resulted in good predictive models.

The objective of this work was to determine whether and to what extent a GMDH network with monthly weather variables as inputs would be able to mimic results of a GCM-based mechanistic simulation of soybean response to climate change and CO_2 elevation.

2. Materials and methods

2.1. Soil and weather data

Three soils typical for the Mississippi Delta were selected for this study. Soil textures were sandy loam, loam, and silt loam (Table 1). The water holding capacities were similar in all three soils.

Data of three global circulation models (GCMs) were used to develop weather data sets for the simulations. Simulation results of the Geophysical Fluid Dynamics Laboratory (GFDL R30) model and of the United Kingdom Meteorological Office (UKMO 89) model were obtained from the National Center for Atmospheric Research (NCAR).[1] The GRDPT utility developed by

[1] ftp://ncardata.ucar.edu/datasets/ds318.1/datafiles/ GFDL ftp://ncardata.ucar.edu/datasets/ds318.2/datafiles/ UKMO

Table 1
Some properties of soils in this study

Depth	Bulk density, g cm^2	Sand, %	Clay, %	Water content at field capacity, cc cm^3 cm^{-3}	Water content at −15 bar, cm^3 cm^{-3}	Saturated hydraulic conductivity, cm day^{-1}
Bosket sandy loam						
26	1.30	53	10	0.381	0.132	3.48
71	1.19	56	10	0.297	0.165	5.29
>71	1.16	75	5	0.362	0.143	5.30
Commerce silt loam						
22	1.18	5	22	0.347	0.288	6.87
81	1.20	4	33	0.362	0.313	1.18
>81	1.17	13	19	0.380	0.319	25.20
Marietta loam						
19	1.42	45	23	0.249	0.189	12.64
49	1.55	27	31	0.279	0.233	15.10
>49	1.37	35	32	0.322	0.266	2.65

NCAR was used to extract twelve monthly values of solar radiation, precipitation and temperature for 1×CO_2 (370 μl l^{-1}) and 2×CO_2 (670 μl l^{-1}) climates for Jackson, Mississippi.

Mechanistic crop simulators use daily weather data whereas GCMs produce monthly averages. We scaled GCM monthly data down to daily weather sequences, using the technique of Haskett et al. (1997) which in turn was based on the method of Wilks (1992). This technique is applied in the DWD (*d*ownscaling *w*eather *d*ata) program. In this technique, the generation of daily weather data from GCM prediction requires (1): parameters of monthly local weather and daily weather variability, and (2) parameters of long-term changes in weather indices averaged over month, season, or year.

Parameters of current baseline weather were estimated for Jackson, MS by Richardson and Wright (1984). Parameters of long-term changes in weather introduced by Wilks (1992) are estimated from monthly weather characteristics produced by GCMs. These parameters are as follow: b_1=the slope of the dependence of the average annual increase in the air temperature on time, δT_{DJF}=standardized average air temperature changes for winter, equal to changes in average winter air temperature divided by change in the global air temperature, δT_{JJA}=standardized average air temperature changes for summer, equal to changes in average summer air temperature divided by change in the global air temperature, ΔR_{DJF}=increment of the average diurnal air temperature range in winter, ΔR_{JJA}=increment of the average diurnal air temperature range in summer, n=exponent relating the increase in the variance of the monthly precipitation to the increase in monthly precipitation, δP_i=relative increases in average monthly precipitation, and δR_i=relative increases in average monthly radiation integral, i=1,12.

We calculate all these parameters (except n, ΔR_{DJF} and ΔR_{JJA}) as differences or ratios from GCM-predicted monthly and seasonal values for the 1×CO_2 and 2×CO_2 years. Examples of the parameters are presented in Table 2. Value n=1.31, ΔR_{DJF}=0, and ΔR_{JJA}=0 were adopted from Haskett et al. (1997).

2.2. Crop simulations

We used the mechanistic soybean crop simulator GLYCIM (Acock and Trent, 1991). GLYCIM consists of a collection of modules that describe related sets of physical or physiological processes and simulate the growth of a plant in a uniform crop that is free of pests and diseases. Balances of materials are kept for individual leaves and petioles on the plant. Balances of materials are kept for other plant organs by type (i.e. stems, flowers, seeds) and for soil by cells. Fluxes of water, heat, nitrate, and oxygen are simulated for the soil while fluxes of carbon, nitrogen, and other structural dry matter are simulated for the plant. This simulator was originally designed to examine interactions between atmospheric CO_2 concentration and other environmental variables acting as factors of soybean crop productivity. An approximately similar level of detail is achieved in descriptions of physical and physiological processes involved in plant growth and development, i.e. light interception, carbon and nitrogen fixation, organ initiation, organ growth and abscission, water and nitrogen flow in soil, heat transport, oxygen transport and intrasoil consumption. The responses of GLYCIM to elevated CO_2 were tested with data from controlled environment chambers and open top chambers (Haskett et al., 1997).

In simulations of soybean yields under climate change, there were four climate samples for each GCM: the climates corresponding to CO_2 concentrations of 370 ppm (considered ambient), 470 ppm, 570 ppm, and 670 ppm. There were 50 replicate weather files for each GCM at each climate sample. A row spacing of 38 cm,

Table 2
Weather generator parameters estimated from two GCM predictions for Jackson, MS

Parameters	UKMO	GFDL
$\Delta T_{2\times CO2}$, °C	4.2	3.9
b_1, °C year^{-1}	0.075	0.06
δT_{DJF}	1.22	1.21
δT_{JJA}	0.86	1.07
Increase in monthly precipitation, %, at doubled atmospheric CO_2 as compared with the current level		
January	21.4	45.9
February	15.9	10.2
March	−4.9	1.0
April	11.4	33.5
May	6.1	4.0
June	25.8	29.9
July	26.5	44.0
August	47.2	41.3
September	26.2	40.6
October	68.7	85.3
November	19.2	19.1
December	1.2	14.2
Relative increase in radiation at doubled atmospheric CO_2 as compared with the current level		
January	1.0	0.96
February	1.0	1.05
March	1.0	1.03
April	1.0	1.01
May	1.1	1.13
June	1.1	1.11
July	1.0	0.98
August	1.0	0.96
September	1.0	0.92
October	0.9	0.86
November	1.0	1.01
December	1.0	0.94

16 plants per meter of a row, 35 kg of residual N–NO_3 per ha, and 3.5 kg residual N–NH_4 per ha, cultivar Pioneer 9592 and emergence on May 15 were assumed in all simulations.

2.3. Group method of data handling

The idea of the Group Method of Data Handling (GMDH) is to employ estimates of the output variable obtained from simple ‘primeval’ regression equations that include small subsets of input variables (Farrow, 1984). Although the accuracy of such estimates is low, it appears that these estimates can be better predictors of the output variable than some of the input variables. The best of these estimates are included in the set of input variables, and again small subsets of variables from this set are used to build new estimates.

A general functioning of the GMDH algorithm can be understood from the following example. Let the original data consist of a column of the observed values of y and of N columns of the observed values of the independent variables $x_1, x_2, \ldots, x_N$. The primeval equations are quadratic polynomials:

$$z=A+Bu+Cv+Du^2+Ev^2+Fuv \quad (2)$$

Here A, B, C, D, E, and F are parameters, u and v are pairs of values of x, and z is the best fit of the dependent variable y.

Each iteration consists of three steps. Step 1 consists of obtaining estimates of y using primeval equations. All independent variables $x_1, x_2, \ldots, x_N$ are taken two at a time to become u and v in Eq. (2), and regression polynomials (2) are constructed to estimate z. The total number of these polynomials is $N(N-1)/2$. The resulting columns of z_m values, m=1, 2, . . ., $N(N-1)/2$, contain estimates of y from each polynomial, and are interpreted as new improved variables that may have better predictive power than the original $x_1, x_2, \ldots, x_N$. The objective of the next step is to keep only the best of these new variables.

Step 2 consists of screening out the least effective new variables. There are several selection criteria (Farrow, 1984) all based on mean square absolute or relative error of values z_m with respect to measured y and often including a correction that ‘punishes’ a network for excessive complexity. In some versions of the method, columns z_m that have the criterion value smaller than a prescribed value are retained. In other versions, a prescribed number of the best z_m are retained. The list of input variables is modified in the end of this step. In some versions of the method, columns of $x_1, x_2, \ldots, x_N$ are merely replaced with the retained columns of $z_1, z_2, \ldots, z_K$, where K is the total number of retained columns. In other versions, the best K retained columns are added to columns $x_1, x_2, \ldots, x_N$ to form a new set of input variables. The total number N of input variables changes to reflect the addition of z_m values or the replacement of old columns x_N with z_m new total number of input variables.

Step 3 consists of testing whether the set of equations can be further improved. The smallest value of the selection criterion obtained at this iteration is compared with the smallest value obtained at the previous iteration. If an improvement is achieved, one goes back and repeats steps 1 and 2, otherwise the iterations stop and the network is built.

For this work, the version of the algorithm coded in the commercial software ModelQuest (AbTech Corp, 1992) was used. It employs three input variables at a time to build primeval regressions as cubic polynomials of three independent variables, utilizes a step-wise regression to diminish if possible both the number of coefficients and the number of independent variables in the primeval equations, limits the number of variables to retain in the input list, and uses the criterion of Barron (1984) to screen out new variables and to stop the iterations when complexity is balanced with accuracy. Both original input variables x_1, x_2, . . ., x_N and output variable y are normalized to have a zero mean value and unit variance value and the normalized variables participate in the network building. Results are then denormalized to the original scale. Since we intended to use GMDH networks as interpolation tools, no attempt was made to divide data sets to training and evaluation subsets.

The original list of input variables for GMDH networks that we developed included average daily values of the radiation integral, maximum and minimum temperature, and precipitation for the five-month period spanning May to September, i.e. growing period for soybeans in the Mississippi Delta. The monthly values were back calculated from daily values generated with the Wilks (1992) method from GCM projections and used in the original simulation. Atmospheric CO_2 concentration was also an input. The total number of original input variables was thus 21, and the GLYCIM-simulated yield was the output predicted by the network. The total number of input–output sets used to build the GMDH network was 400=2 GCMs×4 Climate Samples×50 Replicate Weather Files for each soil.

3. Results

An example of the GMDH network estimating soybean yield on Marietta loam soil is shown in Fig. 1. In this figure, RadAug and RadJul are average daily radiation integral in August and July, respectively, TminJn and TminAu are average minimum daily temperatures in June and August respectively, PrecAu is average August precipitation, CO2 is atmospheric CO_2 concentrations, _Ave and _StdDev mean average and standard deviation across all data sets. Variables x1–x6 are obtained by normalizing RadAug, TminJn, CO2, TminAu, PrecAu, and RadJul, respectively. Yield is obtained by denormalizing output of the network with Yield_Ave and Yield_StdDev which are average and standard deviation of yields across all data sets. Only average daily radiation integral in August and average minimum temperature in June along with the CO_2 concentration were retained at the first iteration, and a new variable z_1 was retained. At the second iteration, average minimum temperature and precipitation in August were selected as additional variables that formed a primeval equation with the z_1 variable to produce the new variable z_2. Finally, at the third iteration, average daily radiation integral in July was selected as an additional variable to form the primeval equation with the z_2 variable. After the third iteration, the number of coefficients in the model was balanced with the model's accuracy and iteration was stopped. Thus only six variables were retained as the essential ones from the original list of 21 variables for this soil. The list of the essential variables appeared to be the same for Marietta loam, and Commerce silt loam soils, whereas the average daily radiation integral in July was replaced with the precipitation in July for the Bosket sandy loam soil.

GLYCIM-simulated and GMDH-estimated yields are compared in Fig. 2. GMDH networks based on monthly weather values explain 81–85% of the variability in yields. Regression of the GMDH-estimated yields on GLYCIM-simulated yields resulted in values of slopes that did not differ significantly from one at 1% probability level.

There were differences between yield data obtained on different soils. The probability distribution functions of yields simulated with GLYCIM are shown in Fig. 3.

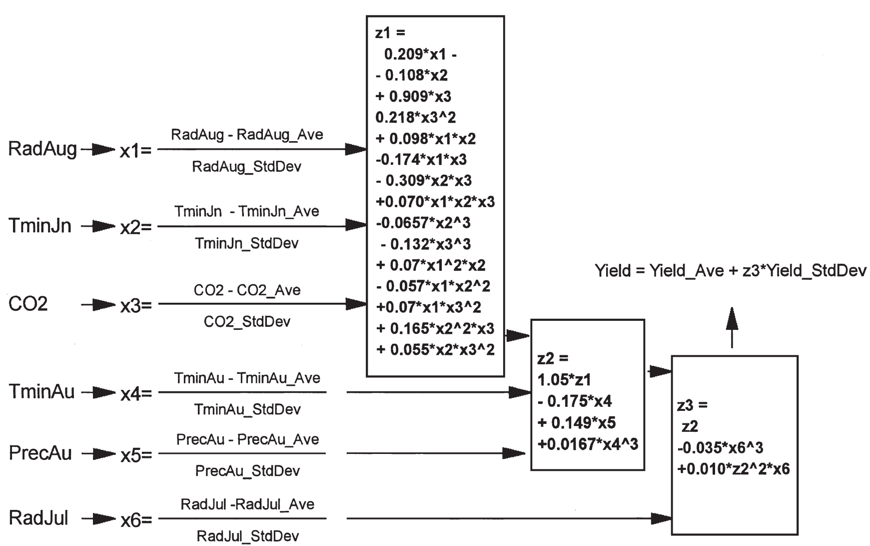

Fig. 1. A GMDH network to estimate soybean yield from monthly weather data and atmospheric CO_2 concentration.

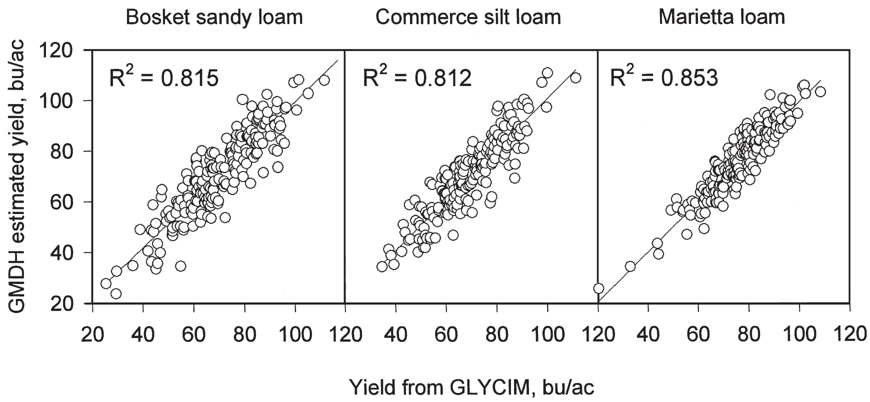

Fig. 2. Comparison of soybean yields calculated with GLYCIM and estimated with the GMDH network for three soils under global change conditions.

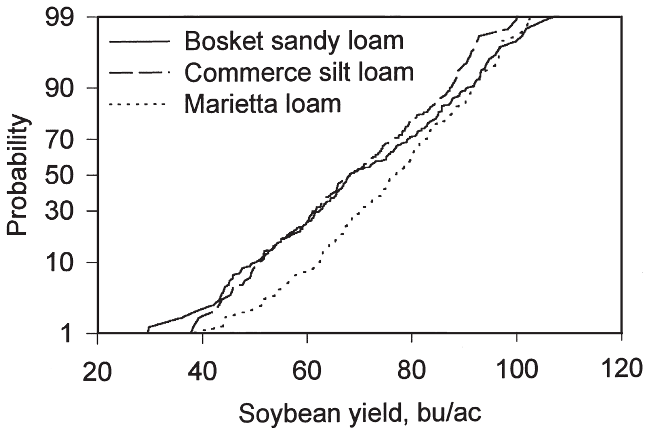

Fig. 3. Probability distribution functions of the soybean yields on three soils estimated under ambient and climate change conditions.

Yields on sandy soil had the widest variation. Yields on loam soil were higher than those on sandy loam and silt loam soil.

A formal sensitivity analysis can be carried out for a GMDH network. However, the transparency of the GMDH dependencies makes it easy to analyze the relative importance of variations in environmental factors for the yield formation. For the example shown in Fig. 1, variable z_3 is approximately equal to variable z_2 in the vicinity of zero values of x_6 and z_2, because both x_6 and z_2 are included only in polynomials of degree higher than one. Therefore, the average radiation integral in July, RadJul represented with the variable x_6 can be important only when it deviates significantly from the average, otherwise the importance of RadJul is only marginal. Analogously, variable z_2 depends on z_1 much more strongly than on x_4 in the vicinity of zero values because coefficients preceding z_1 and x_4 are 1.05 and −0.149, respectively. Therefore the relative effect of precipitation in August reflected in values of x_4 is much less than the relative joint effect of radiation in August, minimum temperature in June, and CO_2 concentration reflected in values of z_1. Finally, the variable z_1 depends on x_3 much more strongly than on x_1 and x_2 because the coefficients preceding x_1, x_2, and x_3 are 0.209, −0.108, and 0.909, respectively. Therefore, the relative effect of the changes in CO_2 elevation on the z_1 values is much larger than that of changes in minimum June temperature or in the August radiation. Going back through the aforementioned chain of reasoning, one can deduce that the CO_2 elevation is the leading factor in yield changes as compared with changes in weather patterns.

Signs of coefficients in GMDH equations also shed light on the effect of the environmental factors. Both x_3 used to calculate z_1 and x_4 used to calculate z_2 have negative coefficients preceding them. The two variables, x_3 and x_4 are normalized minimum daily temperatures, and the yields become lower as the night temperatures increase.

The GMDH dependencies can be used to assess and visualize interactive effects of changes in environmental factors on yields. The effect of changes in CO_2 level and in June minimum temperature is shown in Fig. 4. In this figure, normalized values are obtained by subtraction of average and dividing by the standard deviation, both calculated across all data. As the June minimum temperature increases, the effect of elevated CO_2 on yields becomes mitigated. Yield grows almost linearly with CO_2 increases when June minimum temperatures are low. However, the CO_2 elevation effect on yields becomes mitigated at low CO_2 levels as the June minimum temperature grows. The effect of the June minimum temperature on yields is opposite at low and at high CO_2 levels.

4. Discussion

The GMDH networks were able to reproduce the results of the simulations with a relatively high degree of accuracy. This probably happened because plant development as represented in the CLYCIM simulator is relatively stable and integrates the variable daily

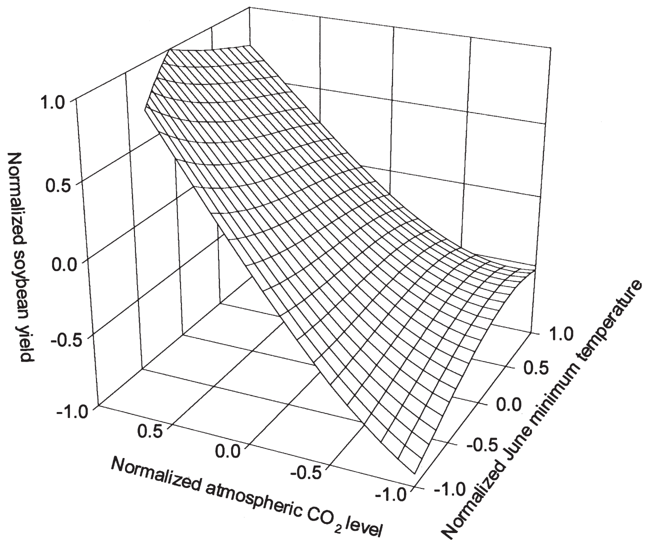

Fig. 4. Sensitivity of the changes in soybean yield on Marietta loam soil to relative changes in atmospheric CO_2 level and in minimum temperature in June.

weather into a relatively robust growth and development pattern. It is consistent with the results of Swanson and Nyankori (1979), who found that the addition of monthly values of precipitation and mean temperature (May, June, July, August) greatly increased the percent of yield variation explained by independent variables in their regression equation. This was particularly true for soybeans for which R^2 increased from 0.285 for the regression with technology alone without weather variables to 0.703 for the regression that included monthly precipitation and mean temperatures. Fluctuations of weather variables within month spans, e.g. dry spells, affect plant growth and development and can limit the accuracy of yield predictions based on monthly average values.

The list of essential monthly weather variables selected by the GMDH is consistent with data from soybean crop research. Researchers agree that water supply is most important for soybean plants during the reproductive development as the plants reach reproductive stages 4 and 5 (Ashley and Ethridge, 1978; Brown et al., 1985; Elmore et al., 1988). August is the month when these stages are achieved by soybean crops in Mississippi (Acock et al., 1997) and this explains the selection of August precipitation as the essential variable.

Both June and August minimum temperatures were selected as the essential predictors. High minimum temperatures result in larger respiration losses of photosynthetically accumulated carbon and also suppress metabolic transformation of carbohydrates to structural compounds in soybean plants (Bunce and Ziska, 1996; Vu et al., 1997). An increase in June minimum temperatures led to increases in yields with ambient CO_2 but caused a decrease in yields with high CO_2 concentrations. High minimum temperatures in June could prevent creation of sufficient leaf area and crop cover to intercept solar radiation in sufficient amounts to utilize extra carbon available for fruit development in August.

The effect of CO_2 elevation on yields was modified by increasing minimum temperatures (Fig. 4). This trend is consistent with results of other researchers who used several GCM scenarios to study effect of climate change on soybean yields. For example, Rosenzweig et al. (1994) did an extensive study on the effects of climate change on crops in all major growing areas of 48 states of the contiguous US. The study used climate data obtained from the GFDL, GISS and UKMO GCMs. The UKMO scenario suggested much larger increases in temperatures than the other two. Effects of changes in atmospheric CO_2 on crop development were included through increases in photosynthesis and stomatal resistance. In this study, CO_2 fertilization mitigated climate-based yield reductions so that yield increases occurred in all but the UKMO scenario in which yields decreased.

CO_2 elevation appears to be the most essential input variable as compared with monthly weather parameters. This result is obtained for a particular crop simulator, and for another crop simulator the relative significance of CO_2 elevation may be different as compared with the significance of the monthly climate variables. Haskett et al. (1997) observed in a study for Iowa that, without the increase in atmospheric CO_2, soybean yields remained practically constant in the GFDL and GISS scenarios, while decreasing by 2 to 10% in the UKMO scenario. In contrast, in a regional study of the effects of climate change and CO_2 elevation on agriculture in the Great Lakes, Ritchie et al. (1989) found that climate change led to decrease in yields in their simulations and resulted in yield increases only in some cases. Such discrepancies show that a simulator needs to be used that is tested with data on plant and crop growth at elevated CO_2.

Soil properties affected the results of simulation in this study. Larger yields were found in loam soil than in the two other soils having extreme textural composition. The generality offered by the GMDH in this study is still limited by the choice of soil for simulations. Though not as weighty as weather, soil properties and plant density remain significant factors affecting crop yield under global change conditions (Haskett et al., 1997). Reddy et al. (1997) found that soil properties had significant effect on carbon partitioning between soybean shoots and roots under climate change conditions in the Southern US. Incorporating the spatial variability of soil and field management into scenarios of simulations and then into inputs of GMDH represents an interesting field to explore.

The results show that GMDH networks can be used to extend the predictive power of computationally intensive mechanistic biological simulation models. This is important for global change studies. Much of the likely behavior of systems of interest in biological global change research occurs outside the range observed in the past. Regression approaches based on historical data are not valid because they involve extrapolations outside the range of the data. Mechanistic simulations based on relevant physical and biological interactions are needed to determine the likely behavior of economically significant species. Interactions between the atmosphere and the biosphere occur gradually over a long period of time. The current work demonstrates that a GMDH network can be used to extend such a dataset that includes intermediate stages in global change progression, to derive general results between crop yields and GCM-based combinations of temperature, precipitation, and CO_2 concentration. This derivation was achieved without prohibitively extensive computation. The need for this kind of flexibility is underscored by the ongoing advances in our understanding of the climate system. These advances drive the evolution of the GCMs resulting in changed climate predictions with each iteration of such models. By deriving the general relationship the likely effects of such new predictions on crop yields can be determined simply by mapping them into GMDH inputs. In this way an initial estimate of the impact of such changes can be obtained which can then be refined if needed by using mechanistic crop simulators.

References

Abdel-Aal, R.E., Elhadidi, M.A., 1995. Modeling and forecasting the daily maximum temperature using abductive machine learning. Weather and Forecasting 10, 310–325.

AbTech Corporation, 1992. ModelQuest. Users Manual. Version 4. Charlottesville, VA

Acock, B., Trent, A., 1991. The soybean crop simulator GLYCIM: Documentation for the modular version. Univ. of Idaho, Agric. Expmt. St. Misc. Series Bull. 145.

Acock, B., Pachepsky, Y., Acock, M., Reddy, V.R., Whisler, F., 1997. Modeling soybean cultivar development rates using data from the Mississippi Valley. Agronomy J. 89, 994–1002.

Alexander, I., Morton, H., 1990. An introduction to neural computing. Chapman and Hall, London.

Ashley, D.A., Ethridge, W.J., 1978. Irrigation effects on vegetative and reproductive development of three soybean cultivars. Agron. J. 70, 467–471.

Barron, A.R., 1984. Predicted squared error: a criterion for automatic model selection. In: Farlow, S.J. (Ed.), Self-organizing methods in modeling: GMDH-type algorithms. Marcel Dekker, New York, pp. 87–104.

Brown, E.A., Caviness, C.E., Brown, D.A., 1985. Response of soybean cultivars to soil moisture deficit. Agron. J. 77, 274–278.

Bunce, J.A., Ziska, L.H., 1996. Responses of respiration to increases in carbon dioxide concentration and temperature in three soybean cultivars. Annals of Botany 77, 507–514.

Elizondo, D.A., McClendon, R.W., Hoogenboom, G., 1994. Neural network models for predicting flowering and physiological maturity of soybeans. Transactions of ASAE 37, 981–988.

Elmore, R.W., Eisenhauer, D.E., Specht, J.E., Williams, J.H., 1988. Soybean yield and yield component response to limited capacity sprinkler irrigation system. J. Prod. Agric. 1, 196–201.

Farrow, S.J., 1984. The GMDH algorithm. In: Farrow, S.J. (Ed.), Self-organizing methods in modeling: GMDH type algorithms. Marcel Dekker, New York, pp. 1–24.

Giorgi, F., Meehl, G.A., Kattenberg, A., Grassl, H., Mitchell, J.F.B., Stouffer, R.J., Tokioka, T., Weaver, A.J., Wigley, T.M.L., 1998. Simulation of regional climate change with global coupled climate models and regional modeling techniques. In: Watson, R.T., Zinyowera, M.C., Moss, R.H. (Eds.), The Regional Impacts of Climate Change. An Assessment of Vulnerabity. Cambridge University Press, Cambridge, New York, pp. 429–437.

Haskett, J.D., Pachepsky, Y.A., Acock, B., 1997. Climate change effects on Iowa soybean production: A comparison of scenarios using GLYCIM. Agronomy J. 89, 167–176.

Hecht-Nielsen, R., 1990. Neurocomputing. Addison-Wesley Publishing Company, Reading, MA.

Jarvis, C.H., Stuart, N., 1996. The sensitivity of a neural network for classifying remotely sensed imagery. Computers and Geosciences 22, 959–968.

Joerding, W.H., Li, Ying, Young, D.L., 1994. Feedforward neural network estimation of a crop yield response function. J. Agric. and Applied Econ. 26, 252–263.

Kleinsteuber, S., Sepehri, N., 1996. A polynomial network modeling approach to a class of large-scale hydraulic systems. Computers Elect. Engng 22, 151–168.

Lein, J.K., 1995. Mapping environmental carrying capacity using an artificial neural network: a first experiment. Land Degradation and Rehabilitation 6, 17–28.

Pachepsky, YA.A., Timlin, D., Varallyay, G., 1996. Artificial neural networks to estimate soil water retention from easily measurable data. Soil Sci. Soc. Am. J. 60, 727–733.

Pattie, D.C., Haas, G., 1996. Forecasting wilderness recreation and use: neural network vs. regression. AI Applications 10, 67–74.

Reddy, V.R., Pachepsky, YA., Marani, A., 1997. Carbon partitioning in cotton and soybean crops in southern US under climate change. World Resour. Rev. 9, 359–371.

Richardson, C.W., Wright, D.A., 1984. WGEN: A model for generating daily weather variables. USDA-ARS.

Ritchie, J.T., Baer, B.D., Chou, T.Y., 1989. Effect of global climate change on agriculture great lakes region. In: Smith, J., Tirpak, D. (Eds). The Potential Effects of Global Climate Change on the US. Appendix C, V. 1., USEPA, EPA-230-05-89-053.

Rosenzweig, C., Hillel, D., 1998. Climate change and global harvest. Oxford University Press, New York, Oxford.

Rosenzweig, C., Curry, B., Ritchie, J.T., Jones, J.W., Chou, T.-Y., Goldberg, R., Iglesias, A., 1994. The effects of potential climate change on simulated grain crops in the United States. In: Implications of Climate Change for International Agriculture: Crop Modeling Study. USEPA, EPA 230-B-94-003.

Sommer, V., Tobias, P., Kohl, D., Sundgren, H., Lundstrom, L., 1995. Neural networks and abductive networks for chemical sensor signals: a case comparison. Sensors and Actuators B 28, 217–222.

Swanson, E.R., Nyankori, J.C., 1979. Influence of weather and technology on corn and soybean yield trends. Agricultural Meteorology 20, 327–342.

Vu, J.C.V., Allen, L.H. Jr, Bowes, G., 1997. Effects of elevated CO_2 and temperature on photosynthesis and Rubisco in rice and soybean. Plant, Cell and Environment 20, 68–76.

Watson, R.T., Zinyowera, M.C., Moss R.H. (eds) 1996. Intergovernmental Panel on Climate Change. Climate Change 1995: Impacts, Adaptations and Mitigation of Climate Change: Scientific-Technical Analyses Contribution of Working Group II to the Second Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, New York.

Wilks, D.S., 1992. Adapting stochastic weather generation algorithms for climate change studies. Climatic Change 22, 67–84.